

Advantages of Increasing Evaluation Capacity in Nonprofits:
How Principles of Process Use Can Inform Development and Strengthen a Nonprofit's Position
in its External Landscape

MPP Professional Paper

In Partial Fulfillment of the Master of Public Policy Degree Requirements
The Hubert H. Humphrey School of Public Affairs
The University of Minnesota

Carley Matsumoto

May 10, 2018

*Signature below of Paper Supervisor certifies successful completion of oral presentation **and** completion of final written version:*

Professor Jodi Sandfort, Paper Supervisor

Date, oral presentation

Date, paper completion

Dean Laura Bloomberg, Second Committee Member

Date

Producer, Jason Matsumoto (Full Spectrum Features)
Third Committee Member

Date

Table of Contents

| | |
|------------------------------------------------|----|
| Introduction | 3 |
| Motivation for this Work | 4 |
| Literature Review | 6 |
| Evaluation Site: Full Spectrum Features | 15 |
| Theory to Practice | 18 |
| Internal/External Analysis | 19 |
| Enhancing Shared Understandings | 21 |
| Infusing Evaluative Thinking | 28 |
| Program and Organizational Development | 32 |
| Conclusion & Discussion | 34 |
| References | 37 |
| Appendices | 40 |

Introduction

Within this paper I explore how process use, or the intentional engagement of stakeholders in the creation and implementation of an evaluation, can increase an organization's capacity for evaluative inquiry. Here, I share the impact of employing process use principles to a specific evaluation conducted within Full Spectrum Features (FSF), a small film production nonprofit based in Chicago, IL.

I begin with a comprehensive literature review of process use theories and practices and continue by exploring the application of relevant principles to an evaluation conducted to assess the effectiveness of FSF's educational tool and short film, *The Orange Story*. This exploration of translating theory to practice focuses primarily on the design phase and initial data collection process; the majority of the data analysis occurred outside the timeline for this paper. Throughout the Theory to Practice section, I incorporate components of the evaluation design, as well as insights from FSF's internal evaluation team, to provide clarity and context. I conclude with a discussion of the challenges and successes of integrating principles of process use into a small nonprofit, then explore the broader implications of how FSF can use evaluation to strengthen their position in external landscapes.

While FSF's capacity for evaluative inquiry will serve as my primary focus, this endeavor cannot ignore the role systematic oppression has played in the erasure of underrepresented stories, in both the past and present. FSF strives to produce materials that have the fortitude to not only insert themselves into the psyche of dominant cultural spaces but to live there, to thrive there and to challenge this space indefinitely. Given this truth, if used intentionally, evaluative inquiry has the potential to increase FSF's legibility in spaces that have historically excluded non-dominant perspectives and ways of knowing.

Motivation for this Work

While much of the work in the nonprofit sector is dedicated to the advancement of marginalized communities, there remains a severe lack of representation in the leadership and financial sponsorship of these endeavors. It is estimated that 90% of nonprofit executives and 85% of board members identify as white (BoardSource, 2017). Individual giving, the third-largest source of nonprofit funding at \$281 billion annually (Giving USA, 2017), is characterized by a donor base that is 73% white (Diversity in Giving, 2015). Further, other common sources of revenue – such as foundation or government grants and contracts – are distributed from institutions that are founded on and guarded by normative ideals of elite whiteness, heteronormativity and able-bodied perspectives.

The hegemony that persists in the nonprofit landscape pressures organizations run by and for marginalized communities to continually package and sell their arguments in a way that is palatable to the institutions who both fund and oppress them. There is a growing awareness of the need to allow individuals from underrepresented populations to own their own stories and lead their own growth. Progress in terms of increasing ownership in this realm, however, has been slow (Guo, Metelsky, Bradshaw, 2012). As one example, the inclusion of individuals who represent the population served on governing boards of nonprofits has challenged the presence of elites, but not in a significant enough way to dismantle elite power structures in these same organizations (Guo et al., 2012).

This is the reality that Full Spectrum Features and peer organizations face on a daily basis. The work produced by FSF strives to advance narratives that have been historically diluted and denied. Filmmakers from diverse communities have used the organization as a haven to explore and expose narratives around race, sexuality, gender and the multitude of intersections

that exist in between. These stories, though fictional, are steeped in the past and present truths of what it means to embody a marginalized identity.

In the context of power and influence, FSF serves as an enclave or a protected site for stories that subvert prevailing narratives to thrive (Mansbridge, 1994). This protected space allows marginalized groups to strategize and energize before they re-enter spaces characterized by rigid homogeneity (Mansbridge, 1994). Evaluation can inform FSF's ability to nurture the diverse perspectives it attracts by enhancing their ability to express truths in a language that is legible in dominant spaces. Here, I address the broader question of how evaluation tools and methods can help organizations like FSF assert marginalized perspectives into dominant spaces.

Internal strength constitutes the first step in this process. Clarifying the objectives of FSF's work and the operations that support them will inform the organization's ability to gather meaningful information to sustain the integrity of their work. Thus, the body of this paper examines how increasing evaluation capacity through process use can inform organizational development and strengthen a sense of collective purpose even within a small nonprofit.

In outlining my motivation for this work, I hope to emphasize that issues of equity and oppression place a constant weight on spaces of color, queerness, disability, etc. Those who operate in these spaces do not have the privilege to ignore these realities or treat them as purely theoretical concerns. I see evaluation and increasing evaluation capacity as an important strategy for FSF to simultaneously safeguard their stories and not only access but thrive within dominant systems. It is thus important to me that the reader maintain this lens as they move through the remaining pages.

Literature Review: Process Use

In a broad sense, I equate the benefits of evaluation to the benefits of practicing meditation. Evaluation provides a means for organizations to reflect internally in order to bring a new awareness to their purpose and practices. This in turn, affects how an organization behaves and is perceived within external environments. In my experience working with small nonprofits, much like meditation, the benefits of evaluation are often known yet the actual practice seems daunting and inaccessible. This is why I chose to focus on transferring skills and knowledge through process use. Process use purposefully engages relevant staff in the practice of designing and implementing evaluations as a means to develop evaluation capacity through practice. It is an inherently integrative and practical strategy which is critical to organizations who lack the means to dedicate time and resources to developing evaluation skills.

The focus on practicality and stakeholder engagement is no surprise given the theorists who have developed and continue to practice process use. In terms of the three branches of evaluation perspectives (valuing, methods and use) the principles of process use were developed by theorists who are part of the use branch (Alkin, 2013). It was first defined by Michael Quinn Patton as the, “individual changes in thinking and behavior, and program or organizational changes in procedures and culture as a result of the learning that occurs during the evaluation process” (1997, p. 90). In other words, stakeholder involvement extends beyond garnering specific insights to inform the design and into the sustained construction of evaluative skills and procedures that can be embedded into the organization’s culture.

Hallie Preskill asserts that, “the concept of process use is based on social constructivist learning theory which suggests that individuals construct knowledge and develop a shared reality through collaboration with others” (2003, p. 427). Within organizations, the facilitation of

process use presents the opportunity for staff to better understand and value evaluation by engaging in evaluation activities that are meaningful to their work (Preskill, 2003, p. 427). Process use falls within the broader concept of evaluation capacity building (ECB), which intends to, “continuously create and sustain overall organization processes that make quality evaluation and its uses routine” (Stockdill, Baizerman, and Compton, 2002). Bradley Cousins describes process use as an indirect approach to ECB as the learning process is embedded in the construction and delivery of an evaluation (2014, p. 9). This is in contrast to direct approaches, which are characterized by an explicit transfer of knowledge that occurs outside the evaluation process (Cousins et al., 2014, p. 9). In short, process use is rooted in the idea that experiential learning can have a sustained impact and encourages continual internal reflection.

An important distinction to be made is the difference between the concepts of process use and findings use. Within program evaluation, the term “use” is commonly employed to describe findings use, or the ability of the organization to make decisions based on the results of an evaluation. While every evaluation should strive to be useful to its primary intended users, the employment of process use is distinct in its focus on gaining skills related to evaluative inquiry (Patton, 2012, p. 142). Marvin Alkin and Sandy Taut distinguish process and findings use through the examination of the two components of learning: 1) knowledge acquisition and accumulation and 2) behavior acquisition and modification (2003, p. 6). While findings use leads primarily to knowledge acquisition and accumulation, process use elicits both types of learning from those involved (Alkin and Taut, 2003, p. 6). Patton echoes and adds to Alkin & Taut’s distinction in his claim that “specific findings typically have a small window of applicability. In contrast, learning to think and act evaluatively can have ongoing relevance” (2012, p. 142).

Now that I have established my working definition of process use, I turn to the strategies and objectives of this practice. An evaluator's employment of process use theories and techniques may have varying goals based on the nature of the organization and program being evaluated, with differing benefits for various stakeholders. That said, increasing internal strength and awareness is a consistent theme among the principles laid out below. Many evaluation theorists have offered insights into the ways process use can impact the individuals and programs within an organization. For clarity, I'll use Patton's six typologies outlined in chapter six of *Essentials of Utilization Focused Evaluation* to reference where theorists and practitioners hold similar or different views (2012). Patton's six typologies of process use include:

- 1. Infusing evaluative thinking into the organizational culture:** According to Patton, this occurs when organizations internalize evaluation processes and incorporate evaluative inquiry in their day-to-day decisions (2012). In other words, evaluation is not treated as an isolated activity, but rather as a critical aspect of the organization's culture. Evaluative inquiry involves challenging and exploring assumptions around to decision making in order to strengthen the actions of an organization (Sharrock, Archibald and Buckley, 2017). Infusing evaluative thinking may take the form of asking evaluative questions to make decisions or building infrastructure that supports evaluation efforts (Patton, 2012, p. 145). James Sanders offers a similar concept which he refers to as "mainstreaming evaluation" (2002, pg. 254).

Forss et al. also highlight the integration of evaluative thinking as a result of process use. They refer to their particular understanding of this idea as "learning to learn" (2002, p. 32). They argue that by engaging in evaluation activities, stakeholders participate in a process characterized by learning and constructing

knowledge. Thus, individuals learn different modes of inquiry that enable them to better understand their programs and activities. (Forss, Rebien & Carlsson, 2002, p. 33).

- 2. Enhancing shared understandings:** A critical aspect of any evaluation is the clarification of the evaluation purpose and intended outcomes. Patton argues that clarifying these components provides a means for staff who hold different perspectives to develop a shared understanding around the organization's objectives. (Patton, 2012, p. 148). Patton asserts that engaging various stakeholders in the evaluation process can help solidify objectives and reconcile existing differences of opinion (Patton, 2012).

While the “creating shared understanding” typology described by Forss, Rebien and Carlsson nearly mimics Patton's wording, there is a subtle difference in their description of the process. Where Patton focuses on creating alignment among stakeholders' varying viewpoints, Forss, Rebien and Carlsson focus on clarifying the different objectives stakeholders may have based on their positions, and on the importance of being explicit as to why those variations are valid (2002, p. 34). In other words, it is perfectly appropriate for individuals to be driven by different objectives given their work's focus. Further, it may be useful to develop a mutual understanding of why these varying motivations exist. These conceptions of “enhancing shared understanding” are not mutually exclusive and can, in fact, complement one another throughout various stages of the evaluation.

- 3. Supporting and reinforcing the program intervention:** Patton claims that evaluation processes can be integrated into a program or intervention in a way that

enhances its activities. *Intervention-Oriented Evaluation*, as he calls it, considers the evaluation to be part of the program, not a separate activity, and seeks creative ways to capture relevant information without disrupting the program's flow (2010, p. 150). Forss, Rebien and Carlsson refer to this concept as "strengthening the project" (2002, p. 36). They assert that activities as simple as interviews conducted by the evaluator or sending a follow-up survey after an intervention is complete can help remind participants of the program or intervention's purpose (Forss, Rebien and Carlsson, 2002, p. 36).

The process of reinforcing and supporting the program can also happen more naturally. Depending on the evaluation project, the staff involved in the process may be exposed to stakeholders they do not interface with on a regular basis. Exposure to the experiences of individuals or groups who hold varying social and professional positions can influence both how staff perceive their program and how they determine the direction of the evaluation (Forss et al., 2002, p. 37).

- 4. Instrumentation effects and reactivity:** An instrumentation effect refers to instances where individuals learn from or react to the measurement tool itself. For instance, Patton explains that the way in which questions are asked on a questionnaire or in an interview can bring awareness to aspects of a program or processes that the participant had not considered (2012, pg.151).

The phrase "what gets measured gets done" captures the reactivity effects of process use (Patton, 2012, p. 151). Essentially, data collection processes can directly or indirectly affect the nature of the work completed by staff. Participating in the evaluation process offers a means for staff to develop guidelines around the intended

outcomes of their work and how to measure success (Patton, 2012, p. 151). Reactivity effects of process use are rooted in the ability of measurements to inform the work that is done.

5. Supporting engagement: Participatory, collaborative, and empowerment

evaluation: Here, Patton discusses the role evaluation can play in reinvigorating key stakeholders' interest in and dedication to a particular initiative (2012, p. 152).

Engaging primary intended users in the process of co-creating an evaluation has the potential to foster ownership among those involved, while garnering useful and relevant perspectives. The opportunity to contribute to an evaluation process provides an inroad for disengaged stakeholders to assert ownership and renew their enthusiasm for their work. The concept of increasing participants' ownership within the evaluation process is the focus of David Fetterman and Abraham Wandersman's work in developing their empowerment evaluation approach. Their approach seeks to engage the community throughout the process, and to impart evaluative tools and skills that can be used for on-going learning and improvement (Fetterman and Wandersman, 2007).

Forss, Rebien and Carlsson echo this notion in their claim that process use can serve to boost morale. According to them, "[the evaluation] reminds [staff] of the purposes they work for and allows them to explore the relationship between their own organization and the development impact that is expected" (Forss, Rebien and Carlsson, 2002, p. 37). Designing and implementing an evaluation demands intentional reflection - a practice that may not naturally occur in many nonprofits but

that can provide the time and space necessary to reconnect with the broader mission and vision of the organization.

6. **Program and organizational development:** Organizational development involves, “making the organization the unit of analysis and organizational effectiveness the focus” (Patton, 2012, p. 158). Analyzing how a program functions within its larger organizational context can lead to improvements that affect the organization as a whole. A program’s effectiveness is certainly influenced by the organizational culture and infrastructure that surrounds it. Highlighting the interconnectedness between program improvements and organizational improvements can lead to ongoing development of both (Patton, 2012, p. 160).

From a programming standpoint, developmental evaluations facilitate learning and program development simultaneously. Patton asserts that, within developmental evaluations, the role of the evaluator as a facilitator is enhanced when all team members engage in the design, data collection, and analysis phases to make informed decisions (Patton, 2012, p. 160).

While the benefits of process use described above can be far-reaching, their impact is largely determined by the attitudes and culture that characterize an organization. Employing principles of process use is not a given; it requires intentional planning on behalf of the evaluator. If we return to the analogy of evaluation as meditation, an individual’s willingness to dedicate time to meditative practices is critical to accruing the benefits. Similarly, key to process use is the evaluator’s belief that a certain site is capable of internalizing evaluative skills and knowledge and will be receptive to the approach. Analysis of the organization’s internal and external environment will help discern whether process use is feasible.

Jean King recommends identifying factors that indicate an organization's willingness or unwillingness to learn through process use (2009, p. 48). Externally, an organization's funding source or critical external stakeholders may influence why and how an evaluation is conducted (King, 2009, p. 48). In other words, it may not be appropriate to employ process use if the evaluation is needed purely to fulfill a grant requirement, or if internal interest in the evaluation is generally low. By contrast, certain environmental opportunities may serve to motivate organizations to develop stronger evaluation capabilities. Growing awareness of issues related to an organization's mission may encourage an organization to prove or improve their outcomes.

Internally, King stresses the importance of assessing the organization's stability, which includes the potential for management support and the organization's general attitudes and experiences toward evaluation (2009, p. 49). Cousins et al. echo King in their claim that the right organizational support structures need to be present for organizational learning to occur. Structural elements that are conducive to process use include low job formalization, which supports ongoing learning of organizational skills and knowledge; communication infrastructure, which enables information to flow horizontally and vertically; and a commitment to professional development (Cousin et al., 2014, p. 15). Without the correct mixture of internal and external characteristics, efforts to employ principles of process use may be in vain.

While process use can be a very useful strategy, it is important to highlight the ways in which it can be misused or potentially damaging to an evaluation's effectiveness. Daniel Stufflebeam asserts that process use, and similar efforts to increase organizational learning, has the potential to detract from designing an effective evaluation, as the evaluator becomes more of an instructor or coach as opposed to a neutral actor (1994, p. 323). A biased relationship where the evaluator focuses too much on the betterment of the organization may lead them to obscure

negative findings in an effort to protect the organization. Data validity can be compromised further if data collection becomes too focused on reflection and skill building as opposed to determining the program's effectiveness (Stufflebeam, 1994, p. 323).

Avoiding these pitfalls largely depends on the evaluator's ability to be transparent about the process and upfront about expectations, particularly around use and misuse of findings. Despite these critiques, I chose to focus on process use as an evaluation capacity building tool for FSF because of its practicality. The evaluation itself provides a means for the staff to reflect on the relationship between their internal functions and values. Process use is critical to sustaining this reflective practice and integrating it into the organization's culture. Building off the idea that FSF serves as a protected site for marginalized individuals, I see process use as an important tool to develop and sustain this space without overburdening the organization's staff and resources.

The Evaluation Site: Full Spectrum Features

Full Spectrum Features (FSF) is a Chicago-based media arts nonprofit. FSF was founded in 2015 and is committed to working with underrepresented communities. Beyond simply diversifying media, FSF seeks to amplify diverse stories by building tools, resources, events, and processes to educate the public on important, but untold histories that have influenced all aspects of society.

FSF is a small organization, with six part-time staff members and several student interns. While the organization's internal staff is small, the number of individuals involved in a single project extends far beyond FSF itself. The creative team of each project, its cast, crew and countless other specialized roles create a complex stakeholder network that ebbs and flows over a project's lifetime. The organization has produced several short narrative films that are rooted in the truths of individuals and communities who are overlooked or intentionally oppressed. While the organization's expertise in elevating these types of stories is well established, FSF continues to explore their ability to extend the life of their work by developing innovative distribution strategies that carve space for this art in both familiar and unfamiliar contexts.

Table 1: Full Spectrum Features Resources and Activities

| Resources | | Activities | |
|----------------------------------------------------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|
| Staff | Unrestricted Funding (top 3 sources) | Films (Active) | Programs |
| Eugene Sun Park, Executive Director Jason Matsumoto, Producer | Grants - (33%) Services- (39%) Unrestricted donations- (11%) | Brujos Signature Movie Make Out Party Formidable Dreams Hidden Histories The Orange Story Chicagoland Shorts | Chicagoland Shorts Expanded Media Illinois Film Tour |

| | | | |
|--------------------|--|----------------------------|--|
| Production Manager | | The T | |
| Program Manager | | Cheetah and the Deathgoers | |
| Designer | | Freelancers Anonymous | |
| Chicagoland Shorts | | Holy Trinity | |
| Lead Programmer | | The Year Between | |
| | | I Am Not Broken | |

This brings us to *The Orange Story*. Many of FSF’s objectives are derived from the production of *The Orange Story*, which is the organization’s first film to intentionally create correspondence between a narrative story and specific learning objectives. *The Orange Story* is set in 1942 and follows a Japanese American man, Koji Oshima, as he abandons his daily life to adhere to Executive Order 9066, which mandated the incarceration of Japanese and Japanese Americans who lived on the West Coast. FSF knew that a short, independently-produced narrative film about Japanese American incarceration during WWII would have a fleeting life and low impact without a distribution strategy that was systematic and far-reaching. After researching ways to sustain the impact of the film, the organization discovered that resources on this topic for educators who teach K-12th grade are particularly scarce (*The Orange Story* Educator Survey, 2016). FSF worked with their Educational Advisory Board to develop a digital education tool to complement the historical moment depicted in *The Orange Story* that would be useful for multiple grade levels.

In this project, the organization aimed to leverage their expertise in narrative filmmaking to develop an empathic learning technique to help students engage with the history of Japanese American incarceration. The concept of empathic learning is based in research on the importance of fostering empathy in classroom settings (Jones et al., 2014). More specifically, *The Orange Story* is rooted in the idea of historical empathy, which fosters a student’s ability to relate to, yet differentiate between, past and present historical contexts, while recognizing that historical

moments continue to hold relevance (Endacott, 2013). The evaluation I conducted in collaboration with a team of internal staff at FSF, centers on the use of both the film and digital education tool in classrooms across the country.

While the primary intended users are the team at FSF, it is important to note that a critical stakeholder in this situation was inaccessible for this particular evaluation. The community of Japanese Americans who were incarcerated characterize the subject of the program technology, but they were not present in the space where evaluative decisions were made. As the evaluator, I had to operate on the assumption that the immense amount of time and effort FSF put into working with this community to create the film and educational tool would enable us to make decisions that respect their reality. While this community is not physically present in the evaluative space, their bodies and voices are incorporated within the digital education tool, which houses primary resources such as oral histories, original press documents and archival photos. It was up to myself and the internal evaluation team to understand how these truths are perceived by our target group. Ultimately, this was an important mindset for me sustain throughout the process given that the evaluation presents the stories of this community as a means to assess students' thoughts and reactions.

Theory to Practice

While the strategies described in the literature review appear fairly intuitive, and a natural part of participating in an evaluation, I found navigating the balance between facilitator and evaluator to be different in practice. The remainder of the paper focuses on the interplay between my ability to facilitate evaluative learning and FSF's capacity to receive and integrate these skills. Table 2 provides an overview of where and when process use strategies were integrated into the evaluation. The concepts listed below represent my assessment of which principles of process use seemed most relevant to the organization in its current state.

Table 2: Relevant Process Use Concepts for Full Spectrum Features

| Process Use Concept | Theorist/Resource | Phase of Evaluation | Activity | Stakeholders involved |
|---------------------------------|------------------------------------------------------|-----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|
| Internal Analysis of FSF | King (2009); Cousins, Goh, Elliott, Bourgeois (2014) | Prior to start of evaluation activities | Evaluator assessment | Evaluator |
| External Analysis of FSF | King (2009); Cousins, Goh, Elliott, Bourgeois (2014) | Prior to start of evaluation activities | Evaluator Assessment | Evaluator |
| Identify Evaluation "Champions" | King (2009) | Prior to start of evaluation activities | Conversations between evaluator and primary intended users | Evaluator; primary intended users |
| Shared Meaning Making | Patton (2012); Forss, Rebien & Carlsson (2002) | Evaluation Design | 1. Preliminary conversations 2. Logic Model development 3. Preliminary Research 4. Purpose statement & question development | Evaluator; primary intended users; internal staff |

| | | | | |
|----------------------------------------|----------------------------------------------------------------|---------------------------------------------------|------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|
| Enhancing Evaluative thinking | Patton (2012); Forss, Rebien & Carlsson (2002); Sanders (2002) | Evaluation Design | 1. Logic Model development, 2. Purpose statement & question development | Evaluator; primary intended users; internal staff |
| Program and Organizational Development | Patton (2012); Forss, Rebien & Carlsson (2002) | Evaluation Design, Data Collection, Data Analysis | 1. Logic Model development 2. Purpose statement & question development 3. Data collection/analysis process | Evaluator; primary intended users; internal staff |

Internal/External Analysis

Intention is an important factor in determining the success of process use. Drawing from King and Cousins et. al's guidelines to determine if process use and/or evaluation capacity building would be viable, I weighed the internal and external factors that might affect Full Spectrum's ability to adopt evaluative inquiry skills (2009, 2014).

Prevailing culture and attitudes around learning and change help to determine whether an organization is equipped to receive evaluative skills (King, 2009). FSF's relative nascency and creative focus both coalesce to create an open, collaborative environment that welcomes exploration. As the organization grows and as its leadership looks inward to advance with intention, I see an opportunity for evaluative inquiry to inform organizational decisions beyond the standalone evaluation for *The Orange Story*. FSF fundamentally understands the need to be intentional about discovering whose story needs to be told and in what context. Given the overarching characteristics of FSF, it was clear there was opportunity for evaluation processes to exist in sustained and meaningful ways.

Once I determined that FSF's culture would be conducive to process use, it was important to identify whether the necessary structures and support were present. Structurally, as the organization is fairly new, a healthy amount of flexibility exists within its operations. The organization is comfortable with adaptation and change, as each film warrants its own creative process and style of collective problem solving. While FSF may not have the structural stability of a more established nonprofit, its culture, flexibility, and intentional engagement in the evaluation process can all inform continued development of necessary structures and support.

After identifying the broader internal and external characteristics, I narrowed my focus to identifying a specific internal change agent. Without an individual who can sustain the knowledge and skills that are transferred throughout this process, evaluation principles may not have as lasting an impact. This calls back to King's claim that it is necessary to identify individuals she describes as "evaluation champions" (2009, p. 48). These are internal staff who have a natural predilection or passion for evaluation *and* have formal power within the organization. Because of FSF's relatively flat structure and small staff, it was fairly easy to determine who could fulfil this role. The necessary characteristics were identified in one of FSF's producers who has worked closely on *The Orange Story*. The producer would serve as the primary contact and leader of the internal evaluation team throughout this process.

Practice implications: I conducted the internal and external analysis in a fairly informal manner. I drew primarily from conversations with the primary intended users about the nature of the work FSF does and how evaluation can inform their success. I also examined the literature on their website related to programming and past endeavors. Despite the informal nature of the analysis, determining FSF's attitudes towards evaluation was a critical first step. In order to

successfully introduce evaluation practices that could be sustained, it was important that FSF not only valued evaluation, but had the capacity to integrate aspects of it into their operations.

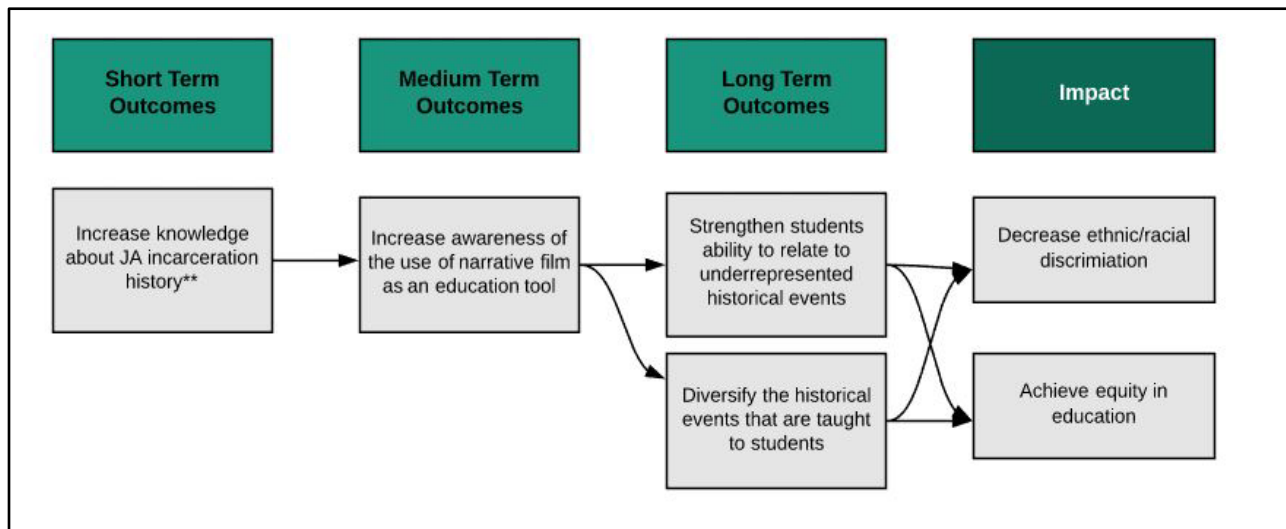
Enhancing Shared Understandings

Enhancing shared understanding was, and continues to be, present through each stage of the evaluation. As the team's concept of how to frame and structure *The Orange Story's* evaluation oscillated between periods of exploration and focus, they developed collective perceptions around how this initiative contributes to the overall organizational objectives. While enhancing shared understanding is an ongoing process that is not confined to any one activity or discussion, there were notable moments where clarity and consensus enabled the evaluation to move forward.

Logic Model Workshop: With few part-time staff and several ongoing projects, building the logic model presented an opportunity for high-level conversation around FSF's broader objectives. The internal team and I had a 2+ hour conversation that began with broad questions around FSF's vision and desired impact and moved backwards through each of the logic model components. Some of the most poignant points of collective understanding surfaced in the team's discussion of how outcomes specific to *The Orange Story* are distinct from, yet nonetheless inform, FSF's ability to achieve its overall objectives. For instance, a specific short-term outcome related to *The Orange Story* is to increase knowledge around the history of Japanese American incarceration during World War II. While this outcome seems unique to this particular film and educational tool, increasing knowledge around Japanese American incarceration directly relates to the organization's ability to achieve its medium-term outcome of increasing awareness of the use of narrative film as an educational tool, and subsequent long-term outcomes

and impact. See Figure 1 for the logic model depiction of how outcomes specific to *The Orange Story* relate to FSF’s overall desired impact of achieving equity in education and decreasing ethnic and racial discrimination.

Figure 1: Full Spectrum Features Logic Model¹



** Indicates outcomes that are specific to *The Orange Story*

The co-conceptualization of FSF’s theory of action enabled us to locate where the evaluation should be situated and enhance the staff’s mutual understanding of how their activities and outputs lead to outcomes. This process also challenged the team to think critically about where their logic is based on assumptions and how their external environment can impact their progress.² Creating a foundation of shared information from which we could build was particularly important given the various levels of familiarity each internal team member had regarding *The Orange Story* and its digital educational tool. Each member’s experiences and

¹ See Appendix A: Full Spectrum Features Logic Model

² See *Challenging Assumptions* section for a more detailed discussion around FSF’s assumptions

perspectives contributed to an environment where no piece of information was assumed or taken for granted.

Determining the Evaluation Purpose and Questions: For us to determine a collective sense of purpose around this evaluation the internal team and I had to reconcile disparate perspectives held by various stakeholders. In this sense, Forss, Rebien and Carlsson's conception of shared meaning making, in which mutual understanding around individuals' ideas is established, was a necessary first step in the process (2002). Our initial conversations around the objectives of the evaluation and how it could inform broader functions within FSF wavered between summative and formative questions depending on each individual's relationship to the project. For instance, the film's director, who also had significant input in the creation of the educational tool, was adamant about understanding how the educational material was being implemented by educators and was primarily concerned with the tool's delivery. While this is important, these questions are specific to *The Orange Story* and wouldn't necessarily lead to broader insights around how FSF's narrative films and educational tools evoke empathy amongst students.

Exploration is, of course, a natural part of creation. The internal team I understood why various stakeholders held different views and we weighed the pros and cons of centering the evaluation around the perspectives we heard. Considering a wide range of possibilities enabled FSF's internal evaluation team to determine what they did and did not want out of this process. Ultimately, we developed a summative evaluation that centered around determining whether *The Orange Story* and the digital education tool evoked empathic responses amongst students.

Table 3: *The Orange Story* Evaluation Purpose, Design & Questions³

| Evaluation Component | Description | Rational |
|------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Evaluation Purpose | Capture the extent to which empathy is present amongst students' responses to <i>The Orange Story</i> and its corresponding digital education tool regardless of their social and political backgrounds | An intended outcome for FSF centers around increasing empathic learning through narrative fiction. This evaluation serves as a first step in understanding if and to what extent empathy is evoked through <i>The Orange Story</i> and its educational tool |
| Evaluation Design | Summative, judgment-oriented evaluation (Patton, 2012) | FSF's intention is to determine whether their work elicits the outcomes it intends. This will be important in determining the overall worth of this distribution method |
| Evaluation Question 1 | To what extent does <i>The Orange Story</i> and its corresponding educational tool evoke empathy amongst students regardless of their experiences? | This particular question seeks to answer whether FSF's learning tool elicits the intended response across students who have diverse life experiences and upbringings. |
| Evaluation Question 2 | How do the self-reported experiences of students relate to the observations made by educators | Collecting information from teachers enables FSF to capture baseline information around whether the empathic responses conveyed by students is typical or not |

There are two interconnected reasons for focusing on this summative approach: 1) *The Orange Story* and education tool serves as a pilot test for FSF and it is critical to the organization to understand whether this model is achieving its desired outcomes; and 2) the creation of a

³ See Appendix E: *The Orange Story* Evaluation Purpose and Design

narrative fiction film *and* a corresponding educational tool are two highly labor-intensive processes.

Aside from the fact that producing high quality fiction requires expensive forms of human capital and equipment, constructing an educational advisory board to design a compelling tool is a separate, labor-intensive process in it of itself. Sharp minds in distinct corners of academia will have to be convinced of the value of FSF's work each time a new model is created. Given these circumstances, the producer and the executive director determined that this particular evaluation should focus on understanding whether the organization is using its energy productively. This process helped solidify FSF's current priorities while acknowledging that different lines of inquiry will have their appropriate time and place in future endeavors.

Challenging Assumptions: As I mentioned previously, the construction of the logic model enabled the internal evaluation team to assess external forces and assumptions that were bigger than themselves. In doing so, the team determined the external factors and assumptions outlined in Figure 2:

Figure 2: Assumptions and External Context⁴

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Assumptions: The combination of empathic learning and diversified curriculum will contribute to a decrease in ethnic/racial discrimination Empathy increases learning outcomes There is a sustained community interest in increasing educational equity and building an inclusive curriculum Students/educators have the resources to access The Orange Story and its digital education tool online | External Context: Funding is currently less stable due to reliance on grants Production of the films/educational tools are dependent on contributions from partners Educator buy-in is essential to the success of the tool |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

⁴ See Appendix A: Full Spectrum Features Logic Model

In particular, the assumption that empathy increases learning outcomes revealed itself to be a particularly important aspect to explore further. If the purpose of the evaluation is to understand if and how empathy appears in the reactions of students, FSF needs to be able to explicitly point to research that supports empathic learning. When I posed the question around what research FSF had done in regards to empathy and empathic learning, the producer disclosed that the organization's understanding drew primarily from conversations with individual educators or education organizations. While these sources may be trusted and legitimate in the eyes of FSF's staff, a more robust body of information would be necessary when communicating with external stakeholders.

FSF's internal team was able to uncover critical research around empathy and educational outcomes that would shape the rest of design process. While the added research prolonged the development of the evaluation design, the exploration of empathy evolved into an exercise that bolstered FSF's internal understanding of how empathy relates to learning. Through this research, we were able to develop data collection tools that would capture information that aligns with pre-existing literature. For instance, the research informed the qualitative coding system we use to analyze student reflections and determine the extent to which empathy is present in their comments. Deepening the internal team's understanding of empathy in the context of learning enabled the team to explicitly connect the work they are doing to concrete concepts and theories as opposed to working off an assumption

Table 4: Empathy Codes⁵

| Code | Definition |
|---------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Affective sharing | Comments that indicate the student felt emotions that mirror the main character's |
| Self-Awareness | Comments that connect the material to the individual's personal life or relevant current events |
| Perspective Taking | Comments in which the student places themselves in the main character's position; This refers to the cognitive ability to learn about the situations affecting others, and to effectively imagine what it would be like to experience the world from the other's position. It requires abstract thought, calculation and applied knowledge. |
| Emotion Regulation | Comments that offer suggestions on how the historical moment or related current events could be handled differently |

Source: (Gerdes, Segal and Leitz, 2010)

Practice Implications: My ultimate intention was to introduce methods that can bring clarity and validity to FSF's work. Throughout the evaluation process, I engaged the internal evaluation team in various modes of inquiry that encouraged them to think collectively not only about the objectives of their work, but how best to support these objectives as well. FSF's ability to ground their work in a foundation of knowledge and logic has already enabled them to both recognize the value of their work internally and legitimize their efforts in external contexts. Since this evaluation's inception, the research the team conducted around empathy has already been used by the *The Orange Story*'s producer in presentations to educators and curriculum developers. Shared meaning making is critical to enabling staff to collect information that informs FSF's development and supports the value of their programs

⁵ See Appendix E: *The Orange Story* Evaluation Purpose & Design

Infusing Evaluative Thinking

Given that we are in the midst of implementing the evaluation, it is difficult for me to assess the extent to which FSF has integrated principles of evaluative inquiry into its culture and infrastructure. Despite this, the staff have expressed invested interest in using various techniques and tools beyond this particular evaluation. FSF is young organization that interfaces with a countless number of innovative and creative individuals. It is integral that the internal staff are equipped to make decisions that are intentional as well as pertinent to the mission and vision of the organization. Similar to enhancing shared understandings, infusing evaluative thinking wasn't confined to any particular activity or conversation and served more as an undercurrent of the evaluation that would surface at various moments. I've highlighted the moments in which evaluative inquiry drove our discussions and presented itself in a way that could be sustained beyond this evaluation.

Developing the evaluation tools: The reason why I include the development of the tools in this section is twofold: 1) the process required the team to continually ask evaluative questions that spanned from usability on behalf of the target group to constructing tools that will capture relevant insights; and 2) the resulting tools and collection system are built so that they can be relevant and useful beyond this evaluation - in other words we built infrastructure that will support future evaluation endeavors (Patton, 2012).

Developing the evaluation tools was a practice in balancing the needs and the capacity of FSF's intended participants (teachers and students) while being able to capture the information that is needed around empathy. One of the primary concerns related to the participants included respecting the educators' time and ensuring our process offered sufficient instruction and resources.

FSF's internal team created a workflow around these core activities that included explicit processes for outreach, communication and the distribution of the necessary materials to each educator who expressed interest. While these activities seem less evaluative and more within the scope of customer service, this practice served to develop evaluation infrastructure that can continue to be used for future initiatives.

In addition to ensuring the evaluation did not inflict an unreasonable amount of burden on the educators, FSF wanted to ensure that the tools minimized the possibility of receiving inflated answers when asking students to self-report on empathy. The internal team's thinking around the data collection tools evolved from a pre and post-test survey that asked students specific questions around empathy to a completely open-ended reflection worksheet refrained from mentioning the concept all together. This change occurred in an effort to avoid asking leading questions. With the open-ended reflection prompts, FSF left room for a range of reactions that could be coded to determine whether aspects of empathy are present.⁶

Table 4: Student Reflection Questions⁷

| | |
|-----------------------|---------------------------------------------------------------------------------------------------|
| Reflection Question 1 | What questions do you have about the content of the film or your assigned chapter? |
| Reflection Question 2 | What observations have you made while watching the film or working through your assigned chapter? |
| Reflection Question 3 | What surprised or did not surprise you about the film or assigned chapter? |
| Reflection Question 4 | What other thoughts or feelings do you have? |

⁶ See Appendix E: *The Orange Story* Evaluation Purpose & Design

⁷ See Appendix C: Student Reflection

The role the educators played in providing insights also evolved as our thinking around the tools progressed. The educators' responses play a critical role in providing a benchmark to which we can compare students' responses. In essence, educator insights can help reveal information around the counterfactual - i.e. what would occur in the absence of the using the film as a teaching tool.⁸

Table 5: Educator Survey Questions⁹

| | |
|-----------------------|-----------------------------------------------------------------------------------------------------------------|
| General Prompt | During the class period, how many of your students... |
| Reflection Question 1 | Connected the material to their personal lives or relevant current events? |
| Reflection Question 2 | Placed themselves in the main character's position (ex. <i>If I were koji, I would have felt...</i>) |
| Reflection Question 3 | Exhibited emotions or displayed an emotional response that mirrored the main characters? |
| Reflection Question 4 | Offered suggestions on how the historical moment or related current events could have been handled differently? |

One notable aspect that is consistent within these tools is the absence of specificity in relation to *The Orange Story* or its educational tool. This again, pivots back to the development of evaluation infrastructure. The internal team did not want the tool to be confined to any one film. Thus, the tools that we co-created can be used to understand how this model evokes empathy across a range of future stories that FSF intends to tell. The transferability of this

⁸ See Appendix E: *The Orange Story* Evaluation Purpose & Design

⁹ See Appendix D: Educator Survey

process in its entirety will be critical as the organization generates new models based off *The Orange Story*. The tools and processes are of course subject to change over time, however, establishing evaluation infrastructure is an important part of accruing relevant longitudinal data that can continue to inform FSF's development.

Developing FSF's logic model: The logic model was one of the few times where a very intentional and explicit transfer of knowledge occurred between myself and the internal team. Because the terminology dictates so much of the process, we could not move forward without ensuring we were all speaking the same language. This was one of the few times I felt like I was instructing as opposed to collaborating and the challenge of structuring the team's thinking to fit within very specific parameters was more pronounced. Despite this, as we moved through each stage of thinking and spent time in each section, it was clear that the internal team benefited from reflecting on the larger implications of their day-to-day actions and how these actions lead to desired outcomes. The type of inquiry the logic model requires resonated with the internal team regardless of the terms we used to frame the discussion.

The logic model itself, in the context of this evaluation has been less of the focal point for the team. While in the future, the artifact may certainly hold relevance as a communication or strategy tool, for this particular endeavor, the logic model encouraged deliberate thinking and the co-conceptualization of FSF's theory of action. While it may be presumptuous to assume this particular type of inquiry will experience sustained use within the organization, it is promising that the three internal team members are equipped to bring evaluative questions to other projects within FSF.

Practice Implications: I believe a critical aspect of evaluative inquiry is its ability to inform both narrow as well as broad processes within an organization. Our examination of each

question within the data collection tools versus our high-level discussion about how FSF's activities lead to outcomes modeled this concept by providing contrasting settings to think systematically about what results the organization can expect and how those results can be achieved (Patton, 2014). For an organization that is fairly young and has the ability to experiment with programming options, this type of thinking can help FSF make informed decisions that better the organization as a whole.

Program and Organizational Development

Process use can inform organizational development by challenging the stakeholders involved to assess a particular program in relation to the organization as a whole (Patton, 2012, p. 152). It is important to note that the producer's professional background includes management and development in the private sector. Thus, this evaluation served to illustrate how this skillset translates to the nonprofit arena. A different set of analytical tools is required to assess organizational outcomes and objectives whose primary intentions are often qualitative or intangible.

The nature of the evaluation around *The Orange Story* places the overall well-being of the organization at the center of the evaluation. FSF staff want to know whether this particular program technology is a viable and sustainable method of bringing histories and stories that are typically overlooked to a broader audience. In other words, this is not an endeavor that is exclusive to *The Orange Story*. Whether or not the film and educational tool achieve its intended outcome has repercussions that extend beyond this particular technology itself - the results will inform future decisions related to FSF's development. The understanding that evaluations, regardless of their scope, are situated in the larger context of the organization's overall impact is

critical for FSF to understand considering their work consists of projects whose focus and audience may span disparate realms.

The most prominent examples related to organizational development via process use have already surfaced within other sections of this paper. The logic model, FSF's continued research around empathy and the development of the purpose statement and questions have all challenged the internal evaluation team to shift between micro and macro lenses to make sense of their work. High level conversations regarding FSF's objectives and opportunities were followed by decisions around word choice and survey heuristics that would elicit empathic responses. Considering that FSF is characterized by staff who are always multitasking on disparate projects, establishing a means of inquiry that can connect these projects to the advancement of FSF as an organization is critical.

Practice Implications: This evaluation around *The Orange Story* is one of the most complex evaluation endeavors the organization has conducted. In a sense, this is a pilot evaluation that can inform future iterations of their work around empathy and narrative fiction. I, nor the staff, can anticipate what the reaction will be from the educators and students. However, whether positive or negative, my hope is that the internal evaluation team is equipped to consider how the results from this specific evaluation can inform the betterment of the organization overall.

Conclusion and Discussion

As an evaluator, this experience has informed how I understand my own role within this field of work. I have garnered skills and insights related to the art of evaluation through both education and practice and that is an immense privilege. With this privilege comes the responsibility (when the situation warrants) to transfer relevant skills and knowledge that can inform an organization's on-going success. This truth is even more pertinent in contexts where issues of equity and oppression both characterize an organization's mission and serve as significant barriers to achieving desired outcomes.

I can't make definitive statements around the extent to which Full Spectrum's capacity for evaluation has increased, however, I can say with confidence that the organization has gained valuable insights into their work. In the current moment, it appears that staff have embraced new modes of inquiry and have, at the very least, been exposed to evaluative tools that are useful beyond the evaluation for *The Orange Story*. However, even if staff claim that evaluation will live a life beyond current endeavors, these statements cannot be verified at this moment. With that said, there are objective indicators that FSF will expand on the processes and skills that have been gained through *The Orange Story* Evaluation.

For one, education on a broader level is written into the mission of the organization. Whether the evaluation around *The Orange Story* yields results that warrant the continued use of this particular distribution strategy, the organization will have learned something about the students and educators they've reached. The infrastructure that has been put in place to capture students' empathic responses to narrative fiction constitute a foundation from which the organization can expand and grow its methods of inquiry. As more feedback is collected and new

patterns that warrant further study emerge, the organization will have the infrastructure to adapt to these needs.

Moving beyond the internal effects of process use, certain parts of FSF's external environment demands more proof than others. This brings us back to FSF's ability to navigate spaces where their work may be unwanted or considered threatening. In the four months I've been working with FSF, their producer has travelled to several educational and curriculum development conferences in Los Angeles, Portland, Seattle and locally in Chicago. While FSF via *The Orange Story* is successfully carving space into its desired landscape, they're excavating exclusively at accessible points of entry. The aforementioned cities and states have common characteristics. Mainly, they're progressive urban centers whose educational systems are likely predisposed to accepting the stories FSF needs to tell. In the meantime, the efforts that have been made to engage regions that are traditionally more conservative have been met with less enthusiasm.

While I do not believe it is the organization's desire to operate in an echo chamber, they currently are not legible in spaces that may view the stories they tell as ones that are rooted in blame or serve as a threat. This evaluation around empathy provides a critical means to reframe how FSF's stories are perceived in such spaces. Outcomes around empathy and the importance of its presence in educational settings shifts the focus away from the fact that *The Orange Story* highlights the prominence of racial oppression in America's past and present and helps FSF communicate the value of their work through a more neutral lens.

It is, of course, not guaranteed that being able to communicate about this particular outcome will grant FSF access wherever they please, but it's the thinking that got them here that might. The nature of FSF's work naturally requires the staff to maintain an awareness of how

internal and external forces shape their success, however, the lens of evaluative inquiry has brought new light to how they can continually capture meaningful information around their work. Within the protected space that FSF safeguards for marginalized voices, evaluative inquiry is a valuable instrument that can inform how the organization presents itself to its desired audience. As FSF learns more about the disparate mentalities that characterize their external landscape, it is my hope that both the tangible and intangible aspects of this process continue to enable the organization to assert untold histories and obscured truths into the spaces that deserve to be challenged.

Resources

- Alkin, Marvin. (2013). *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. SAGE Publications, Inc.
- Alkin, Marvin and Taut, Sandy. (2003). *Unbundling Evaluation Use*,:
<https://www.sciencedirect.com.ezp1.lib.umn.edu/science/article/pii/S0191491X03900010>
- BoardSource (2017). *Leading with Intent: 2017 National Index of Nonprofit Board Practices*. (Pg. 10). Retrieved from: <https://leadingwithintent.org/>
- Cousins, J. Bradley; Goh, C. Swee; Elliott, J. Catherine; Bourgeois, Isabelle (2014). *Framing the Capacity to Do and Use Evaluation*:
<http://onlinelibrary.wiley.com.ezp3.lib.umn.edu/doi/10.1002/ev.20076/full>
- Diversity in Giving: The Changing Landscape of American Philanthropy (2015). Retrieved from:
<http://www.thenonproffitimes.com/wp-content/uploads/2015/03/Diversity-in-Giving-Study-FINAL.pdf>
- Endacott, L. J., (2015). *An Updated Theoretical and Practical Model for Promoting Historical Empathy. Social Studies Research and Practice*.
- Fetterman, David; Wandersman, Abraham. (2007). *Empowerment Evaluation: Yesterday, Today, and Tomorrow*.
- Forss, Kim; Claus, C. Rebien; Carlsson, Jerker. (January 1, 2002). *Process Use of Evaluations: Types of Use that Precede Lessons Learned and Feedback*:
<http://journals.sagepub.com.ezp1.lib.umn.edu/doi/abs/10.1177/1358902002008001515>
- Gerdes, K. E., Segal, E. A., & Lietz, C. A. (2010). *Conceptualising and Measuring Empathy. British Journal of Social Work, 40*(7), 2326-2343. doi:10.1093/bjsw/bcq048

Giving USA (2017). Total Charitable Donations Rise to New High of \$390.05 Billion.

Retrieved from: <https://givingusa.org/tag/giving-usa-2017/>

Guo et al. (2002). Out of the Shadows: Nonprofit Governance Research from Democratic and Critical Perspectives

Jones, S. M., Weissbourd, R., Bouffard, S., & Kahn, J., & Ross, T. (2014). How to build empathy and strengthen your school community. Cambridge, MA: Harvard Graduate School of Education.

King, Jean (2007). Developing Evaluation Capacity Through Process Use.

<http://onlinelibrary.wiley.com.ezp1.lib.umn.edu/doi/10.1002/ev.242/full>

Mansbridge, Jane. (1994). Using Power/Fighting Power. *Constellations*. Vol.1(1), p. 53-73

Mertens, D. M. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of mixed methods research*, 1(3), 212-225.

Preskill, Hallie; Zuckerman, Barbra; Matthews, Bonya. (2003). An Exploratory Study of Process Use: Findings and Implications for Future Research: *American Journal of Evaluation*. Vol.24(4), p.423-442:

<http://journals.sagepub.com.ezp1.lib.umn.edu/doi/abs/10.1177/109821400302400402>

Patton, M. Q. (2012). Identify and Prioritize Primary Intended Uses by Determining Priority Purposes. In *Essentials of Utilization-Focused Evaluation* (pp. 113-139). Los Angeles, California. Sage Publications, Inc.

Patton, M. Q. (2014). Evaluation Flashcards: Embedding Evaluative Thinking in Organizational Culture. *Otto Bremer Foundation*.

The Orange Story Educator Survey, 2016.

Sanders, James R. (2003). Mainstreaming Evaluation, James R. Sanders

<http://onlinelibrary.wiley.com.ezp3.lib.umn.edu/doi/10.1002/ev.89/full>

Sharrock, Guy; Archibald, Tom; Buckley, Jane. (2017). Evaluative thinking as a way of doing business. AEA365 A Tip-A-Day by and for evaluators.

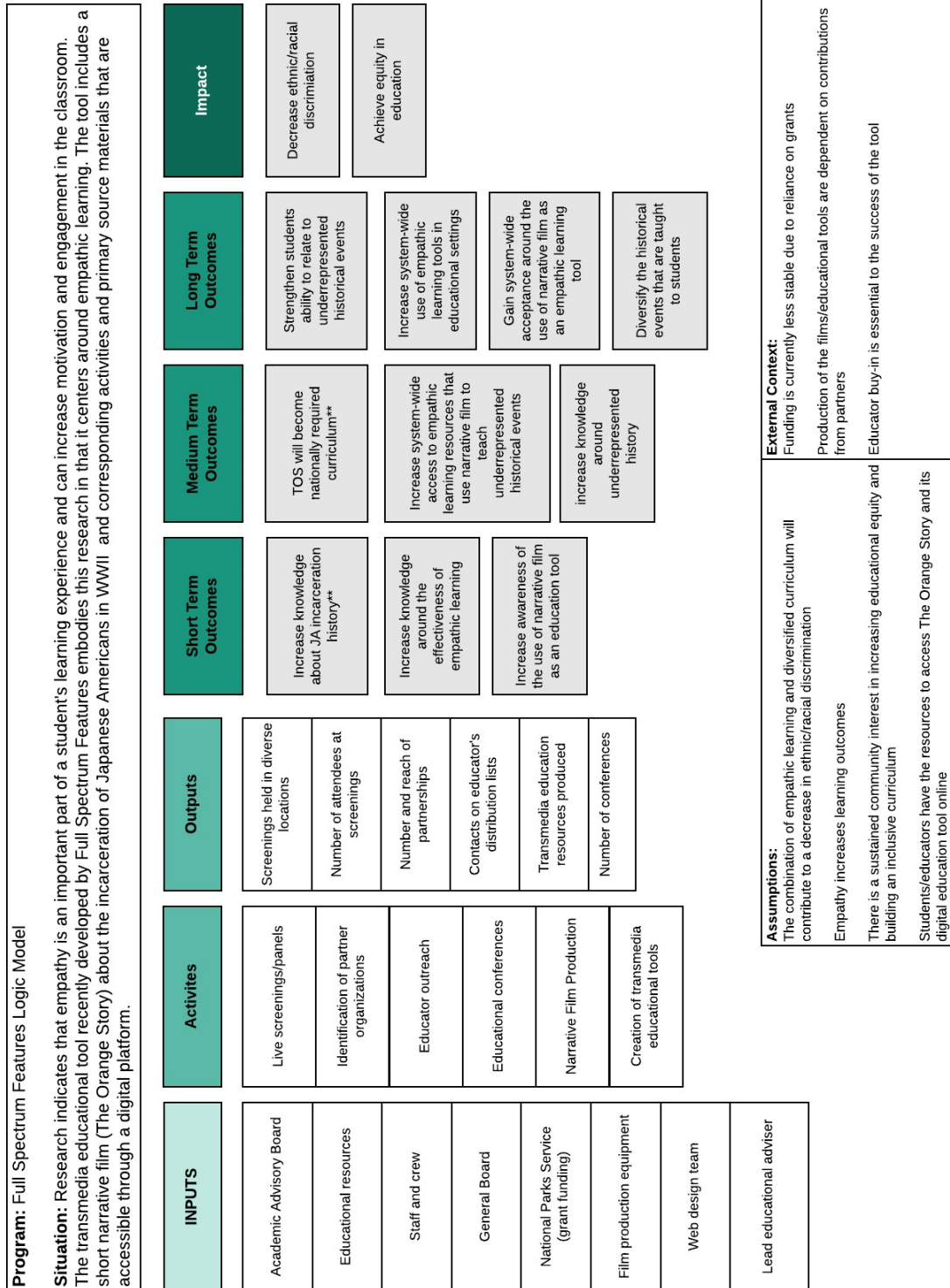
<http://aea365.org/blog/tag/evaluative-thinking/>

Stufflebeam, D. L. (1994). Empowerment evaluation, objectivist evaluation and evaluation standards: Where the future of evaluation should not go and where it needs to go.

Evaluation Practice, 15(3), 321-338.

Appendices

Appendix A: Full Spectrum Features Logic Model



**Outcomes specific to The Orange Story

Appendix B: Instructor Directions



The Orange Story: Directions for Educators

Thank you for agreeing to share *The Orange Story* and the corresponding educational tool with your students. Full Spectrum Features appreciates your willingness to participate in our endeavor and we hope your students enjoy the process. For your reference, deliverables include:

1. STUDENT REFLECTIONS (1 per student)
2. EDUCATOR SURVEY (1 per educator)

Note that you can either choose to submit your deliverables digitally or as hard copies. Details on how to submit each component are provided below.

Directions for Educators:

Preparation

It may be useful to navigate through *The Orange Story's* website to familiarize yourself with the content of each chapter: www.theorangestory.org

Pre-Class Homework Assignment

- Assign all students to watch the full movie (15 minutes):
 - Go to www.theorangestory.org, click on the orange box in the upper righthand corner and choose "Full Movie".
- Assign each student to one of the four chapters of *The Orange Story's* digital education tool with attention to the chapter theme, points of resonance, and the chapter questions ([Ch.1](#), [Ch.2](#), [Ch.3](#), [Ch.4](#)).

In-Class Session

- Watch the film as a class (15 minutes, skip credits)
- Divide students into groups by their assigned chapter for a small-group discussion. Each group will prepare to present the following:
 1. The theme(s) from the assigned chapter;
 2. Three (3) points of resonance;
 3. A collective answer to 1 of the 5 chapter questions
 - a. ([Ch.1](#), [Ch.2](#), [Ch.3](#), [Ch.4](#)).

Appendix C: Student Reflection



The Orange Story Student Reflection

Grade Level _____

Assigned Chapter _____

Directions:

Please navigate to www.theorangestory.com to watch the short film and go through the materials of the chapter you've been assigned

As you watch the film and go through the chapter, please answer the following questions:

- 1) What questions do you have about the content of the film or your assigned chapter?

- 2) What observations have you made while watching the film or working through your assigned chapter?

- 3) What surprised or did not surprise you about the film or assigned chapter?

- 4) What other thoughts or feelings do you have?

Appendix D: Educator Survey



The Orange Story Educator Survey

School Name_____

Grade Level Taught_____

Please circle your answer to the following questions

During the class period, how often did your students...

1. Connect the material to their personal lives or relevant current events
 - ☐ More often than usual
 - ☐ About the same
 - ☐ Less often than usual
 - ☐ Not observed
2. Place themselves in the main character's position (ex. *If I were koji, I would have felt...*)
 - ☐ More often than usual
 - ☐ About the same
 - ☐ Less often than usual
 - ☐ Not observed
3. Exhibit emotions or convey an emotional response that mirrored the main character's emotions
 - ☐ More often than usual
 - ☐ About the same
 - ☐ Less often than usual
 - ☐ Not observed
4. Offer suggestions on how the historical moment or related current events could have been handled differently
 - ☐ More often than usual
 - ☐ About the same
 - ☐ Less often than usual
 - ☐ Not observed

Please continue on to the back of the page...



Qualitative Questions

As an educator please comment on...

☐ Your experience using narrative fiction to introduce the material

☐ Your students' overall reaction to this process

☐ How you would improve the digital tool and/or evaluation process

Appendix E: *The Orange Story* Evaluation Purpose & Design

The Orange Story **Purpose Statement & Evaluation Questions**

Evaluation Context:

Full Spectrum Features (FSF) was founded in 2015 and is committed to both increasing diversity within media arts as well as amplifying the impact of underrepresented artists by creating educational tools that complement their work. These objectives were derived from the production of *The Orange Story*, the organization's first film to intentionally create correspondence between a narrative film and specific learning objectives. *The Orange Story* is set in 1942 and follows a Japanese American man, Koji Oshima, as he abandons his life to adhere to Executive Order 9066.

Full Spectrum Features knew a short, independently-produced narrative film about Japanese American incarceration during WWII would have a fleeting life and low impact without a distribution strategy that was systematic and far-reaching. Upon further research on how to sustain the impact of the film, the organization discovered that resources on this topic for educators who taught within the K-12, middle school and high school ranges are particularly scarce.¹ With this in mind, FSF worked with their Educational Advisory Board to develop a digital education tool that complements the historical moment depicted in *The Orange Story* and is useful for multiple grade levels. Within the context of this project, the organization aimed to leverage their expertise in narrative filmmaking to develop an approach that uses empathic learning as a means for students to engage with the history of Japanese American incarceration during WWII. The concept of empathic learning is based in research around the importance of fostering empathy in classroom settings.² More specifically, *The Orange Story* is rooted in the idea of historical empathy which is defined as a student's ability to relate to, yet differentiate between past and present while recognizing that historical moments continue to hold relevance.³

Purpose Statement:

The purpose of this evaluation is to understand the nature of students' responses to *The Orange Story* and the film's corresponding digital education tool regardless of their social and political backgrounds. The results will be used by FSF to both better understand the value of a model that pairs empathic learning in the form of narrative film with corresponding educational tools as well as communicate more effectively with external stakeholders such as educators and funders.

Evaluation Design:

The principal program technology that FSF has created through *The Orange Story* includes the production of a narrative film that pairs with original digital education materials. This model is innovative in its use of narrative fiction as a foundation to both develop learning materials and introduce students to an historical event. FSF is interested understanding the value of this model by capturing how students resonate with the story.

¹ *The Orange Story* Educator Survey, 2016.

² Jones, S. M., Weissbourd, R., Bouffard, S., & Kahn, J., & Ross, T. (2014). How to build empathy and strengthen your school community. Cambridge, MA: Harvard Graduate School of Education.

³ (Endacott, L. J., (2015). An Updated Theoretical and Practical Model for Promoting Historical Empathy. *Social Studies Research and Practice*.

Given this information, I recommend that FSF conduct a summative, judgment oriented evaluation that captures both the qualitative and quantitative reactions to *The Orange Story* material from students across the country.⁴ In collaboration with FSF's internal staff as well as education consultants, tools that can easily be implemented by educators in their classrooms have been designed. The implementation process is intended to take one class session and includes one survey tool and one qualitative exercise that are integrated into the process of using *The Orange Story* materials. The use of a mixed method approach will enable FSF to gather information that preserves students' unique perceptions of reality while capturing usable information around the tool's impact in relation to empathy that can be communicated to external stakeholders.⁵

Evaluation Questions:

1. To what extent does *The Orange Story* and its corresponding educational tool evoke empathy amongst students regardless of their experiences?
 - a. **Rationale:** FSF's objectives are rooted in research that claims empathy is an important factor in educational settings. This particular question seeks to answer whether FSF's learning tool elicits the intended response across students who have diverse life experiences and upbringings.
2. How do the self-reported experiences of students' relate to the observations made by educators?
 - a. **Rationale:** While the target group for this evaluation are the students themselves, it is important to gather perspectives from those who can offer more objective insights around the students' reactions. By collecting data from both teachers and students we can determine where there is disconnect or alignment between the perceptions of both parties and what that implies in regards to the tool's effectiveness.

⁴ Patton, M. Q. (2012). Identify and Prioritize Primary Intended Uses by Determining Priority Purposes. In *Essentials of Utilization-Focused Evaluation* (pp. 113-139). Los Angeles, California. Sage Publications, Inc.

⁵ Mertens, D. M. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of mixed methods research*, 1(3), 212-225.

Tools & Methods

Overview:

Although historically, the concept of empathy has lacked a single definition and thus resulted in inconsistent measurement methods, literature that includes a review of past definitions as well as more current research related to cognitive neuroscience offers a more comprehensive definition of the characteristics of the term. Gerdes et al. claim that the following four components are necessary to identifying empathy:

1. Affective sharing: This refers to the subjective ‘reflection’ of another person's observable experience (e.g. feeling amused when someone else laughs or sad upon seeing another person cry). This is based on automatic neural mirroring and the ‘shared representations’ such as facial expressions or activities associated with feeling.
2. Self-awareness: Mirroring alone can be so powerful that it effectively erases the perceived boundary between self and other. Self-awareness implies that the empathic person clearly differentiates between his/her experience and that of the person being observed.
3. Mental flexibility and perspective taking: This refers to the cognitive ability to learn about the situations affecting others, and to effectively imagine what it would be like to experience the world from the other's position. It requires abstract thought, calculation and applied knowledge.
4. Emotion regulation: This refers to the empathic person's ability to ‘turn down the volume’ of his/her own feelings as they arise from mirroring another's experience. Inability to regulate emotion can interfere with compassionate action (e.g. by creating overwhelm and burnout for individuals in helping or caretaking social roles).⁶

These components are present in various forms within the research of Mangione et al. whose work continues on to highlight the importance of empathy in learning environments.⁷ In particular, they refer to the concept of *embodied narrative engagement* which refers to the importance of storytelling in “achieving high levels of motivation and engagement in the classroom.”⁸

For the purposes of this evaluation we decided to focus on collecting information around the four components described by Gerdes, et al.: affective sharing, self-awareness, mental flexibility and perspective sharing and emotion regulation.⁹ These components encompass both cognitive and affective forms of empathy which comprise the dual-dimensional conceptualization of the emotional state.¹⁰ It is important to note that this evaluation is focused on determining situational empathy (empathy displayed in reaction to a discrete event) as opposed to dispositional empathy (the extent to which empathy is embedded in an individual's personality).¹¹ That said, we do not aim to assess whether each student

⁶ Gerdes, K. E., Segal, E. A., & Lietz, C. A. (2010). Conceptualising and Measuring Empathy. *British Journal of Social Work*, 40(7), 2326-2343. doi:10.1093/bjsw/bcq048

⁷ Mangione, G. R., Discepolo, T., Tore, P. A., Tore, S. D., Cozzarelli, C., & Corona, F. (2013). Measuring Empathy to Support Learning Design and Narrative Game: A Phenomenological Approach. *2013 Seventh International Conference on Complex, Intelligent, and Software Intensive Systems*. doi:10.1109/cisis.2013.73

⁸ Ibid. pp. 403

⁹ Gerdes, K. E., Segal, E.A., & Lietz, C.A. (2010).

¹⁰ Endacott, L. J., (2015). An Updated Theoretical and Practical Model for Promoting Historical Empathy. *Social Studies Research and Practice*.

¹¹ Eisenberg, N., Fabes, R. A., Murphy, B., Karbon, M., Maszk, P., Smith, M., . . . Suh, K. (1994). The relations of emotionality and regulation to dispositional and situational empathy-related responding. *Journal of Personality and Social Psychology*, 66(4), 776-797. doi:10.1037/0022-3514.66.4.776

possesses all four components, rather we aim to understand how *The Orange Story* evokes empathy amongst students on a collective level.

In order to gather this information, we've designed a mixed-methods approach that gathers the perspectives of both teachers and students. Our objective is to capture the reactions of a wide range of students in order to understand the tool's usefulness in various settings. With this in mind we intend to collect information from at least five classes within the range of 5th -12th grade that are located in distinct regions throughout the country. The following details the implementation plan and tools we've developed:

Implementation Guidelines:

Required commitment: 1 in class session; 1 assignment to be completed outside of class

Implementation process:

- Pre-class assignment
 - Teachers will assign one chapter to each students as well as the Qualitative Feedback Tool for them to complete
 - Students will watch the film in its entirety and go through the resources embedded in the digital education tool of their assigned chapters prior to the class period where TOS will be discussed
- Class time
 - The class will watch the film as a whole (17 minutes, skip credits)
 - Students will break off into small groups based on the chapter they were assigned to discuss three aspects of their chapter that stood out to them (10-15 minutes)
 - Full class discussion where students present on what they discussed (remaining class time)

Tools:

- 1) **Qualitative Feedback Tool:** students will write their answers to the following questions/statements as they go through their chapter:
 - What questions do you have about the content of the film or your assigned chapter?
 - What observations did you make while watching the film or working through your assigned chapter?
 - What surprised/did not surprise you about the content?
 - What other thoughts or feelings do you have?

Rationale

This qualitative method of inquiry is intended to collect information around how students experience Full Spectrum Features' tool. The intention of asking open-ended questions is to elicit responses that can be used to identify whether aspects of empathy are present. In order to avoid influencing students' responses, the questions are intentionally broad and refrain from asking about specific components related to empathy. The results of this exercise have the potential to reveal trends that confirm what staff believe to be true without denying responses that may extend beyond what has been considered thus far.

Analysis

The results will be analyzed using deductive thematic analysis in which we will look for signifiers that relate to the four characteristics of empathy mentioned above: affective sharing, perspective taking, self awareness and emotion regulation.¹² In order to create alignment across both the tool aimed at students as well as the tool aimed at teachers, the system of coding will mirror the questions we ask teachers. In other words we will use the following categories to analyze students' comments and questions:

1. Comments that indicate the student felt emotions that mirror the main character's (affective sharing)
2. Comments that connect the material to the individual's personal life or relevant current events (self-awareness)
3. Comments in which the student places themselves in the main character's position (perspective taking)
4. Comments that offer suggestions on how the historical moment or related current events could be handled differently (emotion regulation)
5. Other

2. Survey Tool for Teachers

Teachers will be provided with a link to the following survey and will be asked to complete it within one day of using the material.

Multiple choice questions around student responses to the material

During the class period, how many of your students...

- Connected the material to their personal lives or relevant current events (self-awareness)
 - More than usual
 - About the same
 - Less than usual
 - Not observed
- Placed themselves in the main character's position. Example: "If I were Koji, I would have..." (cognitive/ perspective taking)
 - More than usual
 - About the same
 - Less than usual
 - Not observed

¹² Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods*, 5(1), 80-92. doi:10.1177/160940690600500107

- Exhibited emotions or convey an emotional response that mirrored the main character's emotions (affective sharing)
 - More than usual
 - About the same
 - Less than usual
 - Not observed

- Offered suggestions on how the historical moment or related current events could have been handled differently (emotion regulation)
 - More than usual
 - About the Same
 - Less than usual
 - Not observed

Qualitative Questions

As an educator please comment on...

1. Your experience using narrative fiction to introduce the material
2. Your students' overall reaction to this process

Rationale:

After conversations with Full Spectrum Features, it was determined that asking students to self report on specific ways the tool affected them may lead to inflated results. Surveying the educators involved will enable us to capture more objective information around students' reactions to the film. It will also enable us to gain broader insights around the tool's impact in the classroom.

Analysis:

Given that our minimum sample size is five classrooms, the use of descriptive statistics to make claims about the tools effectiveness may be limited. However, the educators' responses play a critical role in providing a benchmark to which we can compare students' responses. The quantitative questions are designed to not only establish whether or not a certain type of response was observed but to determine whether responses of that nature occurred more or less than usual. This information will enable us to understand how the tool's ability to elicit specific emotional responses compares to a typical lesson.